# An Introduction to Turbulent Flow

**JEAN MATHIEU**
**JULIAN SCOTT**

Laboratoire de Mécanique des Fluides et d'Acoustique,
Ecole Centrale de Lyon, France

CAMBRIDGE
UNIVERSITY PRESS

PUBLISHED BY THE PRESS SYNDICATE OF THE UNIVERSITY OF CAMBRIDGE
The Pitt Building, Trumpington Street, Cambridge, United Kingdom

CAMBRIDGE UNIVERSITY PRESS
The Edinburgh Building, Cambridge CB2 2RU, UK http://www.cup.cam.ac.uk
40 West 20th Street, New York, NY 10011-4211, USA http://www.cup.org
10 Stamford Road, Oakleigh, Melbourne 3166, Australia
Ruiz de Alarcón 13, 28014 Madrid, Spain



First published 2000

Printed in the United States of America

*Typeface* Sabon 10/12½ pt *System* 3B2 [KW]

*A catalog record for this book is available from the British Library*

*Library of Congress Cataloging-in-Publication Data*
Mathieu, Jean, 1924–
An introduction to turbulent flow / Jean Mathieu, Julian Scott.
p. cm.
Includes bibliographical references.
1. Turbulence I. Scott, Julian, 1954–. II. Title.
QA913.M39 2000 99-16742
532′.0527 – dc21 CIP

ISBN 0 521 57066 2 hardback
ISBN 0 521 77538 8 paperback

# Contents

# Statistical Tools

The apparently random character of turbulent flows strongly suggests that statistical methods will be fruitful. In this chapter, we discuss some statistical techniques, in preparation for later chapters. The reader is presumed to have a modest background in probability theory and this chapter summarizes the main elements, while placing statistical ideas in the context of turbulent flows (those whose knowledge of the basic theory is rusty or nonexistent should perhaps keep one of the many textbooks on the mathematics of probability and statistics to hand; Lumley (1970) and Monin and Yaglom (G 1971) describe the theory as applied to turbulence).

As discussed in Chapter 1, despite being generally considered as deterministic, turbulent flows are highly nonunique in practice. Thus, if an experiment is repeatedly carried out, a different velocity field is obtained in each realization, even if the experimental conditions are nominally the same (i.e., the experimenter endeavors to reproduce the same flow). This is because the detailed behavior of the flow in any one realization is extremely sensitive to small changes in the initial or boundary conditions, which the experimenter cannot control to infinite precision. This type of problem is ideally suited to statistical methods. Indeed, it was the nonrepeatability of an experiment such as tossing a coin that led to the idea of probabilities in the first place. By looking at the statistical properties of an ensemble of different flow realizations, all obtained using the same nominal conditions, one hopes to extract useful quantities, namely probabilities and averages, which depend only on parameters that the experimenter controls. To take an example, the detailed behavior of the turbulent wake of a sphere placed in a uniform flow may vary with tiny perturbations in the incoming stream and small vibrations of the sphere, to name just two possible extraneous factors, but one hopes and expects that, for instance, the average velocity is well defined. Of course, the flow statistics depend on the gross experimental conditions used, for example, a turbulent boundary layer can have quite different thickness and drag with suction than without, but we expect them not to change significantly with small variations in those conditions.

Given an ensemble of different flow realizations, we can define the associated probabilities of flow variables[1] taking on particular values, or more precisely, ranges of values, since they are generally continuous variables. Thus, we imagine an experiment performed very many times under nominally the same conditions, each time producing a different realization of the flow, and use the frequency with which a

[1] Flow variables range from simple ones, such as the pressure at a single point and time, to more complicated ones that are tensorial or obtained from the flow at many points and times.

given flow variable falls into a given range of values as a definition of the probability for that range.

Mean (or average) values are of particular importance in the theory of turbulence. We can define means by taking the average over an increasingly large number of realizations of the flow under the same nominal conditions. The mean value can also be calculated in the usual way from a knowledge of the probabilities of the given quantity as a sum (or integral for continuous variables) over all possible values, weighted by their probabilities. In particular, the mean flow is defined as having the average values of the velocity and pressure (together with the density and perhaps other fluid properties if the fluid is compressible, a case which is not considered in this book). The departure of any given realization from the mean can be calculated by subtracting the mean flow and is conventionally identified with turbulence. That is, the total flow is split into a mean part and a fluctuating component, whose average is zero and which is usually thought of as representing the turbulence. For instance, the mean values of the squared fluctuating velocities are often used to characterize the intensity of turbulence. We shall come across numerous other important quantities defined by averages during the course of this book.

## 2.1 Probabilities and Averaging

Consider any flow variable $U$, which might represent a velocity component or the pressure at a given position and time, or a more complicated quantity derived from different points and times. $U$ will take on different values in different experiments. One may average the quantity $U$ by summing over experimental realizations, dividing by the number of realizations, and letting that number go to infinity. Thus we obtain the mean of $U$, variously denoted as $\overline{U}$ or $\langle U \rangle$ (or sometimes $E(U)$, then called the expectation of $U$). The probability distribution function, $P(U)$, can also be defined so that

$$\int_{U_-}^{U_+} P(U)\, dU \tag{2.1}$$

gives the proportion of realizations in the ensemble for which $U$ takes on values in the range $U_- < U < U_+$, that is, the probability that it falls in that range. The idea here is that $P(U)dU$ gives the probability of $U$ lying between $U$ and $U + dU$ and that we should sum up such elementary contributions over the range $U_-$ to $U_+$ to determine the overall probability that $U_- < U < U_+$. The probability of a given event is thus the proportion of the ensemble for which it occurs and (2.1) gives the probability that $U_- < U < U_+$.

Since $U$ must always take on some value,

$$\int_{-\infty}^{+\infty} P(U)\, dU = 1 \tag{2.2}$$

is an identity which the distribution function, $P(U)$, must always satisfy. The mean of $U$ can also be calculated from $P(U)$ via

$$\overline{U} = \int_{-\infty}^{+\infty} UP(U)\, dU \tag{2.3}$$

which expresses the fact that the average can be computed by taking the proportion of the sample in the range $U$ to $U + dU$, which is $P(U)dU$, multiply it by $U$ and sum over all possible values of $U$. We can likewise determine the mean of any function of $U$ from

$$\overline{f(U)} = \int_{-\infty}^{+\infty} f(U)P(U)\, dU \tag{2.4}$$

Given $U_- < U_+$, if we take $f(U)$ as the window function

$$\begin{aligned} f(U) &= 1 \qquad \text{if} \qquad U_- < U < U_+ \\ f(U) &= 0 \qquad \text{if} \qquad U \leq U_- \text{ or } U \geq U_+ \end{aligned} \tag{2.5}$$

the average value of $f(U)$ is

$$\overline{f(U)} = \int_{U_-}^{U_+} P(U)dU \tag{2.6}$$

that is, the probability that $U$ falls in the range $U_- < U < U_+$, from which $P(U)$ can be obtained by differentiation with respect to $U_-$ or $U_+$. Thus, knowledge of the probability distribution function allows calculation of average values from (2.4), while, working in the other direction, averaging can be used to derive the distribution function. Because mean values are often easier to determine experimentally, averaging of window functions has formed the basis of a popular experimental technique for the measurement of probability distributions. In this method, an experimental signal, $U(t)$, is sent through an electronic device whose output is 1 for a certain range of values and 0 outside that range, followed by averaging to obtain the probability that $U(t)$ lies in the given range. Experimentally, time averaging is usually employed for steady flows, whereas we have defined both probabilities and mean values via an ensemble of experiments. Conditions under which the two approaches give the same results will be discussed in the next section.

An obvious property of averaging is that it is a linear operation, that is, if $\lambda$ is any constant

$$\overline{\lambda U} = \lambda \overline{U} \tag{2.7}$$

and, if $V$ is any flow quantity

$$\overline{U + V} = \overline{U} + \overline{V} \tag{2.8}$$

Since flow quantities are usually functions of spatial location and time (e.g., a velocity component), governed by differential equations, we will frequently need to take averages of their derivatives. Linearity allows one to write

$$\frac{\overline{U(x+h) - U(x)}}{h} = \frac{\overline{U}(x+h) - \overline{U}(x)}{h} \tag{2.9}$$

and taking the limit as $h \to 0$ we have

$$\overline{\frac{\partial U}{\partial x}} = \frac{\partial \overline{U}}{\partial x} \tag{2.10}$$

which indicates that one can take averages inside derivatives, an operation we will often perform in later chapters without explicitly noting the fact. Note that this holds for both space and time derivatives and that similar results apply to integrals over either space or time. Mathematically, we might say that averaging commutes with differentiation and integration.

Most quantities occurring in the theory of turbulence are continuous variables, that is, they take on a continuum of values and the value in one experiment is never exactly the same as that in any other. This behavior may be contrasted with that of discrete variables (e.g., the number of times the velocity exceeds a certain critical value in a given time range), which can only take on values within a certain countable set of numbers and thus tend to repeat themselves. The distribution function of a discrete variable consists of Dirac functions, whose amplitudes give the associated probabilities. For instance, the trivial case in which the variable $U$ takes the same value $U_0$ in all realizations is described by $P(U) = \delta(U - U_0)$, where the amplitude is 1 since that gives the probability that $U = U_0$. One can also consider variables which are part continuous and part discrete, whose distribution functions consist of Dirac functions embedded within a non-Dirac continuum. Notice that, although the probability distributions of continuous variables derived from turbulence are usually smooth functions, in general the probability distribution of a continuous variable may be a discontinuous function. For example, a uniformly distributed random variable in the range $0 < U < 1$ has $P(U) = 1$ in that range, and zero outside. However, we do not intend to open the door on the various mathematical pathologies which might arise and, at worst, the distribution functions we have in mind for continuous variables show isolated jump discontinuities of the type illustrated by this example.

Examples of continuous $U(t)$ and their probability distributions are shown in Figures 2.1–2.3. In interpreting these figures, we assume that mean values can be calculated using time averaging, detailed conditions for which are given in the next section. The first example, Figure 2.1, is typical of a turbulent quantity in a steady flow (statistical steadiness being one of the conditions for use of time averaging). The figure shows the time history, $U(t)$, the probability distribution of $U$, and the result of conversion of $U(t)$ to 0 and 1 values, with 1 when $U$ lies in some range $U_- < U < U_+$, an operation that might, in practice, be performed by an electronic gate. As discussed above, the mean value of the resulting signal should give the probability that $U$ lies in the given range. This mean, and hence the probability that $U_- < U < U_+$, can be obtained by time averaging the output of the gate. The distribution function,

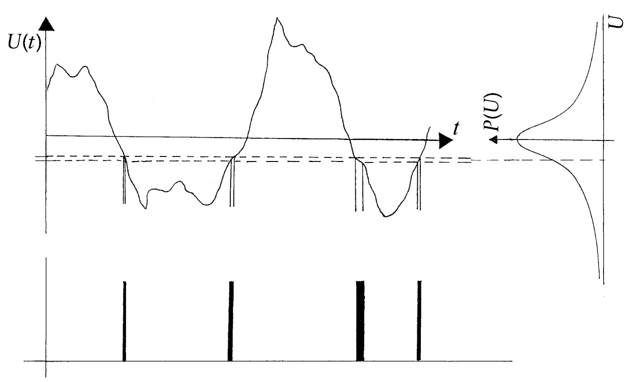


**Figure 2.1.** Sketch of the time history of a typical turbulent random quantity with its probability distribution function. Also shown is the result of passing the signal through a device that produces an output of 1 if the quantity lies in a given range and 0 if it lies outside the range. The average of the device output is the probability that the quantity lies in the range.

$P(U)$, perhaps determined in the above manner, is shown in the figure and has a single hump, typical of many turbulent flow quantities. The tails of the distribution, which represent rare values of $U$, are the most difficult to measure accurately, since such values only occur infrequently and so very long sampling times are needed to obtain converged statistics. Figure 2.2 illustrates a less common type with a double-humped, strongly asymmetric distribution. Figure 2.3 shows a sine wave, which might come from a nonrandom periodic flow if the amplitude and phase of the wave are repeatable from one experiment to another. In that case, $U$ takes the same value in all realizations and is not truly random (nor statistically steady, so one cannot use time averaging to obtain mean values). It is then a discrete variable with only one possible value whose probability is 1, yielding a distribution function consisting of a single Dirac function, whose position is a sinusoidal function of time. If, on the other hand, and as implicit in the figure, the phase of the wave varies randomly from one realization to another with uniform probability over the range 0 to $\pi$, then the signal is statistically steady, with the probability distribution shown in the figure and mean value equal to zero. One could also allow the phase to vary between realizations over a different range of values or nonuniformly, in general yielding statistics that vary periodically with time. At first sight, such sine waves appear to have little to do with turbulence, but if one imagines adding a certain amount of the signal in Figure 2.1 to that in Figure 2.3, the result resembles a periodically modulated turbulent flow, such as that produced by blowing across the mouth of a bottle to produce a tone. A turbulent shear layer is produced over the mouth of the bottle, which, coupled to the Helmholtz resonance of the bottle, yields a turbulent flow with nearly periodic, self-sustained oscillations. If the phase of the oscillations varies uncontrollably from realization to realization of the flow, as is likely unless the oscillations are somehow phase locked, the oscillations themselves will appear as random fluctuations, whereas if the phase is the same in all realizations, they appear as a periodic mean flow superimposed on a periodically modulated random component due to turbulence in the shear layer. We will return to this example in the next section.

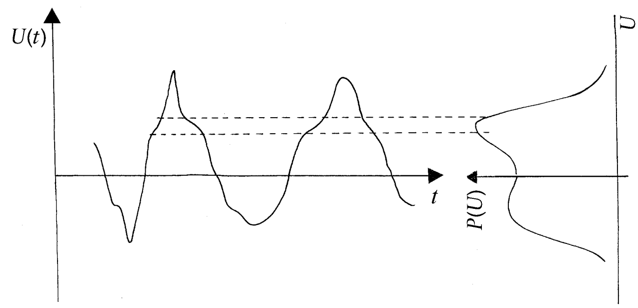


Figure 2.2. Time history and distribution function of a random quantity having a double-hump probability distribution.

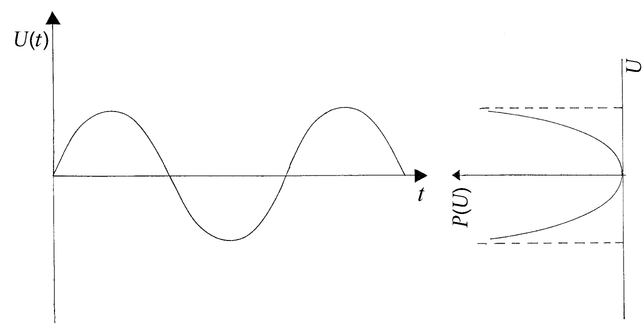


Figure 2.3. Time history of a single realization and probability distribution for an ensemble of randomly phased sinusoids.

Given several flow quantities, one can define a joint probability distribution function. For example, in the case of two variables, $U_1$ and $U_2$, the proportion of the ensemble in which $U_1$ takes values between $U_1$ and $U_1 + dU_1$, and $U_2$ takes values between $U_2$ and $U_2 + dU_2$ is given by $P(U_1, U_2)dU_1dU_2$. Thus,

$$\int_{U_{2-}}^{U_{2+}} \int_{U_{1-}}^{U_{1+}} P(U_1, U_2)\, dU_1 dU_2 \tag{2.11}$$

gives the probability that $U_{1-} < U_1 < U_{1+}$ and $U_{2-} < U_2 < U_{2+}$. Clearly

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} P(U_1, U_2)\, dU_1 dU_2 = 1 \tag{2.12}$$

and

$$P_1(U_1) = \int_{-\infty}^{+\infty} P(U_1, U_2)\, dU_2 \tag{2.13}$$

$$P_2(U_2) = \int_{-\infty}^{+\infty} P(U_1, U_2)\, dU_1 \tag{2.14}$$

give the distribution functions of $U_1$ and $U_2$ individually. The average of any function $f(U_1, U_2)$ can be obtained as

$$\overline{f(U_1, U_2)} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(U_1, U_2) P(U_1, U_2)\, dU_1 dU_2 \tag{2.15}$$

The quantity $P(U_1, U_2)$ is often referred to as the joint distribution of $U_1$ and $U_2$, while the expression "joint statistics" is also sometimes used. Note that nothing stops us taking $U_1$ and $U_2$ as the values of a single flow quantity at different fixed times, $t_1$ and $t_2$, or at different spatial locations. It should be apparent how the above ideas can be extended to an arbitrary number of flow variables. The joint probability distribution of $N$ variables, $U_1, \ldots, U_N$, is a function of those variables and can be thought of as a scalar field in the $N$-dimensional space with coordinates $U_1, \ldots, U_N$. For instance, the probability distribution, $P(\mathbf{U})$, of the vector velocity, $\mathbf{U} = (U_1, U_2, U_3)$, at a given point and time in a turbulent flow is a scalar function in the three-dimensional space spanned by the vector $\mathbf{U}$. Thus, the probability that the vector $\mathbf{U}$ falls within a small volume element, $d^3\mathbf{U}$, of that space is $P(\mathbf{U})d^3\mathbf{U}$.

Conditional probabilities and averages may be introduced as follows. The basic idea is to restrict attention to those realizations in which some flow quantity, $U_2$ say, takes on a particular value, that is, experiments in which $U_2$ has values other than the given one are ignored. Within this subensemble, the proportion of experiments in which $U_1$ has values between $U_1$ and $U_1 + dU_1$ is denoted by $P(U_1|U_2)dU_1$. Thus, $P(U_1|U_2)$ is the probability distribution of $U_1$, conditional on $U_2$ having the given value. When $U_2$ is a discrete value of the random variable on which the statistics are conditioned, this definition is satisfactory, but, as it stands, it does not work otherwise, because $U_2$ will never have the given value exactly. In that case, one allows $U_2$ to take on a small range of values between $U_2$ and $U_2 + dU_2$, and defines $P(U_1|U_2)$ as the proportion of such experiments that also give $U_1$ in the range $U_1$ and $U_1 + dU_1$. Now, the proportion of the total ensemble satisfying both conditions is $P(U_1, U_2)dU_1dU_2$, while the proportion of the total in which $U_2$ lies in the range $U_2$ to $U_2 + dU_2$ is $P_2(U_2)dU_2$. It follows, after a little thought, that

$$P(U_1|U_2) = \frac{P(U_1, U_2)}{P_2(U_2)} \tag{2.16}$$

which can be taken as the definition of the probability distribution of $U_1$, conditional on a nondiscrete value of $U_2$. Thus, given the joint distribution function, $P(U_1, U_2)$, of continuous variables, one can calculate the conditional distribution functions from (2.14) and (2.16). Notice that there is a difficulty when $P_2(U_2) = 0$, mathematically apparent through division by zero. Suppose, for instance, that $P_2(U_2) = 0$ over some range of values. It is evident that there is little sense in trying to define statistics conditional on values of $U_2$ in that range, since they never occur. Thus, $P(U_1|U_2)$ is undefined when $U_2$ is a continuous value of the conditioning variable and $P_2(U_2) = 0$. From (2.13) and (2.16), we obtain

$$P_1(U_1) = \int_{-\infty}^{\infty} P(U_1|U_2)P_2(U_2)\,dU_2 \tag{2.17}$$

showing that the single-variable distribution function can be calculated by combining the conditional distributions for all possible values of the conditioning variable, weighted by their probabilities. One may extend the above ideas to the joint distribution of any number of random variables with multiple conditioning variables.

Conditional averages can also be defined. Thus, the mean of $U_1$, taken only over realizations in which $U_2$ has a given value, is

$$\langle U_1|U_2\rangle = \int_{-\infty}^{+\infty} U_1 P(U_1|U_2)\,dU_1 \tag{2.18}$$

and can be used to construct the unconditional average via

$$\overline{U_1} = \int_{-\infty}^{\infty} \langle U_1 \mid U_2\rangle P_2(U_2)\,dU_2 \tag{2.19}$$

This result shows that one may calculate the average of $U_1$ in two stages. First determine the conditional averages with $U_2$ fixed, then average of all possible values of $U_2$. This may seem like a rather indirect way of proceeding, but such an approach sometimes proves the easiest way of determining average values. Once again, extension to multiple conditioning variables is straightforward.

Conditional averages are often useful in interpreting data from turbulent flows. Consider, for instance, the example of a boundary layer shown in Figure 2.4. The frontier of turbulence is sharp and mobile. Sometimes a given point finds itself inside the turbulence and sometimes it is outside. Data obtained inside and outside will be quite different in character and, if one simply takes the average of some flow quantity, the result does not generally reflect what is happening in either region, but instead gives some intermediate value. However, provided one can experimentally identify when a given sensor lies inside and outside the turbulent region, two conditional averages may be calculated: the first giving the average value inside the turbulence, the second outside. These conditional values will provide more detailed information than the unconditional average and the technique can be applied whenever one suspects statistics that differ significantly under identifiably different circumstances. In the case of the boundary layer, the variable, $U_2$, on which the averages are made conditional is discrete. Thus, although the frontier of turbulence is not really infinitely thin, a threshold value for some measure of turbulence intensity

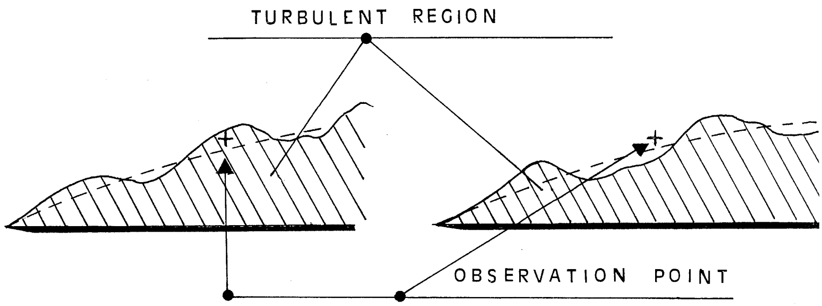


**Figure 2.4.** Sketch of a turbulent boundary layer at two different times, illustrating the use of conditional averaging. The hatched region represents the zone of turbulence. The dashed line is the mean location of the frontier between turbulent and laminar flow, while the continuous line is its instantaneous location.

is employed and one declares the measurement point to be inside the turbulence if it exceeds the threshold. The results should be insensitive to the choice of threshold, since the frontier of turbulence is thin.

Conditional averaging is often implemented by gating the signal, $U_1(t)$, whose conditional average one wishes to measure. Suppose, for instance, that we wanted to determine the average of some flow variable $U_1$ at a point in a boundary layer, conditional on being inside the turbulence. Let $U_2 = 0$ outside the turbulence and $U_2 = 1$ inside, denoting the probability of the latter by $p$. Thus $p$ is the probability of turbulence at the given point. Gating simply takes $U_1$ and $U_2$ as inputs and outputs $U_1$ if $U_2 = 1$, and zero otherwise. The average of the gate output can be calculated as follows. In a large number, $N$, of realizations, the number in which the output is nonzero is $pN$. The average of the output over those nonzero realizations is $\langle U_1 \mid U_2 = 1\rangle$, so the total sum of the different outputs over all $N$ realizations is $pN\langle U_1 \mid U_2 = 1\rangle$. Dividing by $N$, we obtain the average output

$$\overline{U_1 U_2} = p\langle U_1 \mid U_2 = 1\rangle \tag{2.20}$$

while, as discussed earlier, $p = \overline{U_2}$. Thus, we can calculate the conditional average as

$$\langle U_1 \mid U_2 = 1\rangle = \frac{\overline{U_1 U_2}}{\overline{U_2}} \tag{2.21}$$

This result indicates that we should divide the averaged gate output by the probability that the given measurement point lies inside the turbulence, which can be obtained by averaging $U_2$, to determine the required conditional average. The probability that a given point lies inside the turbulence is often called the turbulent intermittence and, in the case of the boundary layer, decreases with distance from the body surface owing to the decreasing frequency of turbulence, from a value very close to 1 in the region near the wall, to 0 outside the layer. The increasing rarity of turbulence at larger distances means that longer time samples are needed to obtain convergence of the averages in (2.21). Given that one can measure conditional mean values, conditional probabilities can be obtained by conditionally averaging functions which are 1 inside some range and 0 elsewhere, as for unconditional probability distributions. As the range used becomes narrower, one again needs to sample for longer to obtain converged statistics.

From (2.16), if $U_1$ and $U_2$ are statistically independent of one another, that is, specifying $U_2$ does not affect the distribution of $U_1$ and vice versa, we have

$$P_1(U_1) = P(U_1|U_2) = \frac{P(U_1, U_2)}{P_2(U_2)} \tag{2.22}$$

or

$$P(U_1, U_2) = P_1(U_1)P_2(U_2) \tag{2.23}$$

which expresses the important fact that the joint distribution function of independent variables is simply the product of their individual distribution functions. One consequence of (2.23) is that

$$\begin{aligned}\overline{U_1 U_2} &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} U_1 U_2 P_1(U_1)P_2(U_2)\,dU_1 dU_2 \\ &= \int_{-\infty}^{+\infty} U_1 P_1(U_1)\,dU_1 \int_{-\infty}^{+\infty} U_2 P_2(U_2)\,dU_2 = \overline{U_1}\,\overline{U_2}\end{aligned} \tag{2.24}$$

showing that the mean of a product of statistically independent factors is the product of their means. In many turbulent flows, it is found that the velocity at widely separated times appears to approach statistical independence as the temporal separation increases. That is, if $U(t_1)$ and $U(t_2)$ are velocity components at the two times, they approach independence as $|t_1 - t_2| \to \infty$. Put another way, knowledge of the flow at the earlier time does not tell us much about its later behavior, which can be expressed by saying that the flow has only limited statistical memory. We will return to this topic in the next section when discussing correlation functions.

## 2.2 Statistical Moments and Correlations

Given a quantity $U$, the mean value of any power,

$$a_\nu = \overline{U^\nu} = \int_{-\infty}^{+\infty} U^\nu P(U)\,dU \tag{2.25}$$

is called the moment of order $\nu$ or $\nu$th moment of $U$. The moment of order 1 is, of course, the mean of $U$. The central moments are defined by

$$\mu_\nu = \overline{(U - \overline{U})^\nu} = \int_{-\infty}^{+\infty} (U - \overline{U})^\nu P(U)\,dU \tag{2.26}$$

and, apart from $\mu_1$, which is zero, are generally more important than the $a_\nu$ in the theory of turbulence. Distribution functions exist for which the integrals in (2.25) and (2.26) fail to converge for some values of $\nu$, which means that the corresponding moments do not exist. However, this is not usually the case for variables derived from turbulent flows, at least not for the positive orders of moments we have in mind here. The most important central moment is obtained for $\nu = 2$ and is called the variance, given by

$$\sigma^2 = \mathrm{Var}(U) = \mu_2 = \overline{(U - \overline{U})^2} \tag{2.27}$$

where $\sigma \geq 0$ is the standard deviation of $U$ and measures how far about its mean $U$ varies, that is, the magnitude of the random fluctuations in $U$. Furthermore, if $U_1$ and $U_2$ are statistically *independent*, it is not difficult, using (2.24), to derive

$$\mathrm{Var}(U_1 + U_2) = \mathrm{Var}(U_1) + \mathrm{Var}(U_2) \tag{2.28}$$

which can be extended to show that the variance of the sum of any number of independent variables is the sum of their variances.

The next two central moments can be nondimensionalized using $\sigma$ to obtain the skewness

$$S = \frac{\mu_3}{\sigma^3} = \frac{\overline{(U - \overline{U})^3}}{\sigma^3} \tag{2.29}$$

and flatness factor (or kurtosis)

$$T = \frac{\mu_4}{\sigma^4} = \frac{\overline{(U - \overline{U})^4}}{\sigma^4} \tag{2.30}$$

of $U$. These higher-order moments are of considerably less importance than the variance, but the skewness is one possible, if coarse, measure of lack of symmetry of the distribution of $U$ about its mean, whereas the flatness factor provides limited information about how extensive the tails of the distribution are.

Given two variables, $U_1$ and $U_2$, we can define their correlation

$$R = \overline{(U_1 - \overline{U_1})(U_2 - \overline{U_2})} \tag{2.31}$$

and, expressing the same quantity in nondimensional form, the correlation coefficient

$$\rho = \frac{R}{\sigma_1 \sigma_2} = \frac{\overline{(U_1 - \overline{U_1})(U_2 - \overline{U_2})}}{\sigma_1 \sigma_2} \tag{2.32}$$

which always lies between $-1$ and $+1$ and is zero for statistically independent variables. To derive the bounds on $\rho$, let $u_1 = U_1 - \overline{U_1}$, $u_2 = U_2 - \overline{U_2}$ be the fluctuations in $U_1$ and $U_2$, so that $R = \overline{u_1 u_2}$. Since

$$\overline{(u_1 + \lambda u_2)^2} \geq 0 \tag{2.33}$$

for any constant $\lambda$, expanding the square leads to

$$\sigma_2^2 \lambda^2 + 2R\lambda + \sigma_1^2 \geq 0 \tag{2.34}$$

for any $\lambda$, which implies that

$$|R| \leq \sigma_1 \sigma_2 \tag{2.35}$$

or, in other words, $|\rho| \leq 1$, as stated above. Observe that if $\rho$ takes on one of its limiting values, $\rho = \pm 1$, then the above argument shows that $\sigma_2 u_1 = \pm \sigma_1 u_2$ and the random variables are deterministically related, since their fluctuations have the same ratio in all realizations. To show that $R = \rho = 0$ when the variables are statistically independent, we note that

$$R = \overline{u_1 u_2} = \overline{(U_1 - \overline{U_1})(U_2 - \overline{U_2})} = \overline{(U_1 - \overline{U_1})}\,\overline{(U_2 - \overline{U_2})} = 0 \tag{2.36}$$

where we have used (2.24) to write the mean of the product as a product of the means. It follows that $R$ and $\rho$ are measures of statistical dependence, although it should be noted that they might be zero even if the variables are not statistically independent.

As remarked at the end of the last section, many turbulent flows are thought to asymptotically approach statistical independence at wide temporal separations. This is reflected in correlation coefficients that go to zero at large temporal separation, though not necessarily monotonically. Thus, if $U(t)$ is some velocity component, $U_1 = U(t_1)$ and $U_2 = U(t_2)$ decorrelate as $|t_1 - t_2| \to \infty$, that is, the correlation coefficient $\rho(t_1, t_2) \to 0$. At zero time separation, $\rho(t_1, t_2) = 1$ takes its maximum value, from which it falls away at nonzero $|t_1 - t_2|$. The order of magnitude, $\Theta$, of the temporal separation required for significant decorrelation is referred to as the correlation time. Decorrelation is also often observed between the velocities at two points in space at a single time as a function of spatial separation. The distance required for significant spatial decorrelation, or correlation length, is an important measure of the size of the large scales of turbulence, as discussed in Chapter 1. Chapter 3 examines the different time and space scales present in turbulent flows in some detail, but for the moment we want to consider the process of time averaging of flow quantities.

We noted earlier that time averaging is often used for experimental determination of the statistical properties of steady flows and we now want to pose the question as to when this leads to the same results as the ensemble definition. The flow is assumed statistically steady, for otherwise time averaging will mix together differing statistics from different times and there is no real hope that it will yield ensemble averages corresponding to a specific time. Consider some time-dependent flow quantity $U(t)$ and define the time average as

$$\hat{U}^{(T)} = \frac{1}{T}\int_0^T U(t)\,dt \tag{2.37}$$

whose ensemble average yields

$$\overline{\hat{U}^{(T)}} = \frac{1}{T}\int_0^T \overline{U}\,dt = \overline{U} \tag{2.38}$$

since the flow is supposed statistically steady. This shows that the ensemble average of $\hat{U}^{(T)}$ agrees with the ensemble average of $U$. Subtracting (2.38) from (2.37), squaring and ensemble averaging, we find

$$\mathrm{Var}\left(\hat{U}^{(T)}\right) = \overline{\left(\frac{1}{T}\int_0^T u\,dt\right)^2} = \frac{1}{T^2}\int_0^T\int_0^T \overline{u(t_1)u(t_2)}\,dt_1 dt_2 \tag{2.39}$$

where $u(t) = U(t) - \overline{U}$ is the fluctuation in $U$. Introducing the correlation function

$$\overline{u(t_1)u(t_2)} = R(t_1 - t_2) \tag{2.40}$$

which is a function of the temporal separation $t_1 - t_2$ alone, thanks to statistical steadiness of $U(t)$. Changing integration variables to $t_1$ and $\tau = t_1 - t_2$, instead of $t_1$ and $t_2$, the integral over $t_1$ can be performed to give

$$\mathrm{Var}\left(\hat{U}^{(T)}\right) = \frac{\sigma^2}{T}\int_{-T}^{T}\left(1 - \frac{|\tau|}{T}\right)\rho(\tau)\,d\tau \tag{2.41}$$

where $\rho(\tau) = R(\tau)/\sigma^2$ is the correlation coefficient at time separation $\tau$ and $\sigma$ is the standard deviation of $U$. Suppose that $\rho(\tau) \to 0$ sufficiently rapidly as $|\tau| \to \infty$ that the integral

$$\Theta = \int_{-\infty}^{\infty} |\rho(\tau)|\,d\tau \tag{2.42}$$

converges, yielding a correlation time.[2] One can then bound the integral in (2.41) to obtain

$$\sigma\left(\hat{U}^{(T)}\right) \leq \left(\frac{\Theta}{T}\right)^{1/2}\sigma \tag{2.43}$$

where $\sigma(\hat{U}^{(T)})$ is the standard deviation of $\hat{U}^{(T)}$. It follows that $\sigma(\hat{U}^{(T)}) \to 0$ as $T \to \infty$, that is, the random fluctuations of the finite time average, $\hat{U}^{(T)}$, about its mean value, $\overline{U}$, can be made as small as one likes by increasing the averaging time, $T$. That is, if $T$ is taken large enough, the fluctuations in $\hat{U}^{(T)}$ are very small and the time average is a good estimate of the ensemble average. We conclude that time averaging over a sufficiently long period yields the same results as ensemble averaging for steady flows provided that there is rapid enough decorrelation that (2.42) converges. In that case, equation (2.43) can be used to estimate the error involved in the time average. The error decreases only slowly with increasing averaging time, proportional to $T^{-1/2}$. This type of calculation is often made when designing statistical measurements of turbulent flows. One asks how long an averaging time is needed for convergence of the average to within an acceptable margin of error. As we saw earlier, the determination of probability distributions can be reduced to the calculation of appropriate averages.

The time average (2.37) is inappropriate for unsteady flows; for instance, in the case of turbulent flow generated by an explosion, one must repeatedly carry out the experiment to produce the ensemble statistics. If a time average is employed it includes all stages of the explosion and neither converges nor yields results which meaningfully describe any given stage. However, consider the example sketched in Figure 2.5 of a cylinder inside a piston engine that is turning at constant speed and load. In such a flow, there are turbulent fluctuations from one cycle to the next, so a single realization of the flow is not periodic, but we might expect the statistical properties of the flow, for instance the mean velocity, to vary periodically with time. That is, the flow statistics vary throughout the cycle, but those at time $t$ are the same as those at time $t + \tau$, where $\tau$ is the period of oscillation of the piston. In that case, a time average can be defined by

$$\hat{U}^{(N)}(t) = \frac{1}{N}\sum_{n=0}^{N-1} U(t + n\tau) \tag{2.44}$$

[2] Although precise expressions, such as (2.42), for correlation times arise in particular circumstances, in general it is better to think of them as order of magnitude scales. The same is true of correlation lengths.

rather than (2.37) and an argument very similar to that used above shows that, provided $N$ is sufficiently large and there is rapid decorrelation of $U(t + n\tau)$ with increasing temporal separations, (2.44) will yield a good approximation to the ensemble average. As before, probabilities can be determined by appropriate averaging.

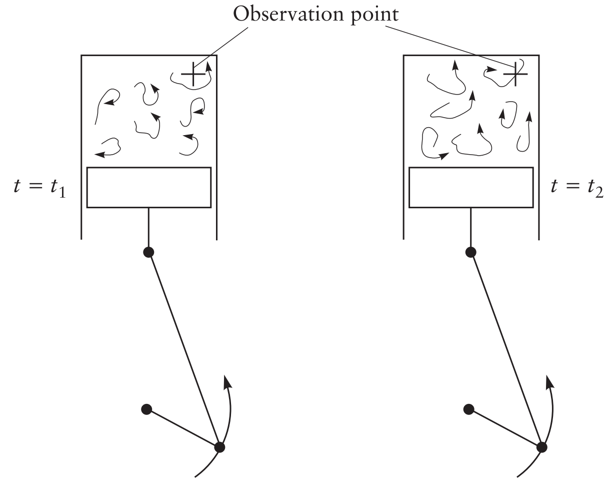


**Figure 2.5.** Illustration of a periodic turbulent flow. The two times shown, $t_1$ and $t_2$, are an integral number of periods apart so that the flow is nominally the same. It will, in fact, be different in detail because of turbulent fluctuations, as we have attempted to show schematically. The turbulent-velocity fluctuations sketched are superimposed on a periodic mean flow, which is the same at the two times.

Blowing across the mouth of a bottle at high Reynolds number to produce an audible tone provides an example in which, if one had never done the experiment, one might expect the flow to be nonoscillatory, whereas, in fact, it has important, nearly periodic oscillations, together with fluctuations from cycle to cycle of the oscillations due to turbulence in the periodically modulated jet/shear-layer over the mouth of the bottle. In this example, the flow itself generates self-sustained oscillations, in contrast with that of Figure 2.5, whose periodic variations are externally imposed by the piston. Repetition of the experiment to provide an ensemble yields oscillations whose phase varies from realization to realization, as in the earlier example of the randomly phased sine wave and reflecting the fact that the experimenter does not control the phase. As with the sine wave, the statistics may turn out to be independent of time, owing to random phasing, but this does not accurately reflect our intuition about the flow. The periodic oscillations, which are not really random in the intuitive sense, are lumped in with the turbulence as part of the random fluctuations and, in consequence, correlations extend to large temporal separations (in principle, to infinite separations if the oscillations were precisely periodic, although exact periodicity is unlikely in practice, given the possibility of random phase drifting over many cycles). This is an unsatisfactory situation because one would like to separate the physically distinct fluctuations due to the oscillations from those occurring from cycle to cycle, which one might identify with turbulence. One way of doing this is feasible if the phase of the oscillations can be experimentally identified in particular realizations, for the flow statistics may then be *conditioned by restricting attention to the subensemble of realizations in which the phase has a particular value* (or even by time-shifting the data so that the phase becomes the same in all realizations). The pressure fluctuations at some point within the bottle provide a good measure of the oscillations because they strongly focus attention on the Helmholtz resonance, rather than on the turbulent fluctuations in the jet. Thus, if the statistics are conditioned by the internal pressure fluctuation, we expect the oscillations to appear as a periodically varying mean flow, while periodically modulated fluctuations about the mean reflect turbulence within the jet and may now decorrelate rapidly with temporal separation. One can then use phase-locked sampling to calculate averages from (2.44). Conditioning the statistics in the above manner effectively makes the phase a controlled parameter and the flow becomes fundamentally similar to that of Figure 2.5, having statistics which vary periodically with time, more closely expressing our intuition about the nature of the flow. We should also remark that Fourier analysis in time allows one to identify periodic, or nearly periodic, components of a flow, which appear as sharp peaks in the Fourier transform, corresponding to the oscillation period and its harmonics.

The problem of uncontrolled parameters, like the phase in the above example, appears in spades when one considers naturally occurring flows, such as the atmospheric boundary layer in which we live. Furthermore, other than in the imagination, one cannot repeat the experiment to generate an ensemble of realizations, unless laboratory or numerical simulations of sufficient fidelity can be constructed, a possibility we ignore for the sake of argument. Thus, we are reduced to passive spectators, although the flow can be observed at different times and locations. In the case of the atmospheric boundary layer, Fourier analysis of the wind velocity with respect to time shows that its transform has well-separated peaks corresponding to time scales of the order of one minute and four days, which represent boundary-layer turbulence and the passage of meteorological systems. It is thus reasonable to employ time averages of the velocity with an averaging time large compared to the smaller of these scales and small compared with the larger, allowing statistical properties of boundary-layer turbulence to be measured. This procedure makes the longer time scales part of the average, while the shorter ones become fluctuations, hopefully decorrelating at temporal separations larger than a few minutes, if not to zero, then at least to small values. The layer statistics depend on the wind speed above the layer and its thermal stratification, to name but two time-varying parameters that the experimenter has no control over. In a theoretical model, one might consider an ensemble of realizations in which all such parameters are held fixed, presumably avoiding the difficulties associated with uncontrolled parameters and multiple time scales for variation of the real flow. Provided the parameters do not vary too rapidly in reality, one would expect such ensemble statistics to agree with those measured using appropriate conditioning or time averaging.

In summary, when ensembles of realizations are used to define the statistics of turbulence they should not be defined blindly, but with the physical properties of the flow in mind. In particular, all important parameters of the flow ought to be fixed, otherwise one may end up with fluctuations that include components of the flow other than turbulence and correlations that extend over large time or space separations.[3] This being said, most fundamental studies of turbulence concern more straightforward cases than those envisaged above, for instance, simple jets and boundary layers at high Reynolds numbers, or turbulence generated by passage of a uniform, steady flow through a grid. For these relatively simple flows, in which the experimenter is presumed to control all important parameters, it suffices to consider a full ensemble of realizations generated by repeating the experiment, or equivalently, since the flows are usually steady and decorrelate with temporal separation,

[3] To caricature, it has jokingly been said that, once one has eliminated all features of a flow that one understands, what remains is turbulence.

time-averaged statistics. One hopes and expects that experience gained with such flows will extend, at least in part, to turbulence in more complicated situations, such as atmospheric and oceanic flows.

We now turn to more mundane matters, namely the definition of characteristic functions and cumulants. The significance of cumulants lies in the fact that those of order higher than two are zero for Gaussian variables, an important class of statistics we discuss in the next section, but, for the moment, the definition of cumulants is purely formal.

The first characteristic function, $\varphi(s)$, of a random variable $U$ is the complex-valued quantity given by

$$\varphi(s) = \overline{e^{isU}} = \overline{\cos sU} + i\,\overline{\sin sU} = \int_{-\infty}^{+\infty} e^{isU} P(U)\, dU \tag{2.45}$$

which will be recognized as the Fourier transform of the distribution function $P(U)$. Some mathematical properties of $\varphi(s)$ are as follows. Given $\varphi(s)$, we can determine $P(U)$ by Fourier inversion, so that $\varphi(s)$ contains the same information as the distribution of $U$. It is easily seen that $\varphi(0) = 1$ and $\varphi(-s) = \varphi^*(s)$, where "$*$" denotes complex conjugation, while if $P(U) = P(-U)$ then $\varphi(s)$ is real and $\varphi(-s) = \varphi(s)$. Writing the exponential in (2.45) as a power series, we obtain

$$\varphi(s) = \sum_{n=0}^{\infty} a_n \frac{(is)^n}{n!} \tag{2.46}$$

where $a_n$ are the moments of $U$, while a similar procedure using $e^{isU} = e^{is\overline{U}} e^{is(U-\overline{U})}$ leads to

$$\varphi(s) = e^{is\overline{U}} \sum_{n=0}^{\infty} \mu_n \frac{(is)^n}{n!} \tag{2.47}$$

where $\mu_n$ are the central moments. The second characteristic function is defined by

$$\Psi(s) = \log \varphi(s) \tag{2.48}$$

where a principal value for the logarithm is implied. The power series expansion of $\Psi(s)$ is

$$\Psi(s) = \log\left\{ e^{is\overline{U}} \sum_{n=0}^{\infty} \mu_n \frac{(is)^n}{n!} \right\} = \sum_{n=1}^{\infty} \kappa_n \frac{(is)^n}{n!} \tag{2.49}$$

where $\kappa_n$ are referred to as the cumulants of $U$. After some algebra, it can be shown that

$$\begin{aligned}
\kappa_1 &= \overline{U} \\
\kappa_2 &= \mu_2 = \overline{u^2} = \sigma^2 \\
\kappa_3 &= \mu_3 = \overline{u^3} \\
\kappa_4 &= \mu_4 - 3\mu_2^2 = \overline{u^4} - 3\overline{u^2}^2 \\
\kappa_5 &= \mu_5 - 10\mu_2\mu_3 = \overline{u^5} - 10\overline{u^2}\,\overline{u^3} \\
\kappa_6 &= \mu_6 - 10\mu_3^2 - 15\mu_2\mu_4 + 30\mu_2^3 = \overline{u^6} - 10\overline{u^3}^2 - 15\overline{u^2}\,\overline{u^4} + 30\overline{u^2}^3
\end{aligned} \tag{2.50}$$

The cumulants are mainly significant because they are zero, apart from $\kappa_1$ and $\kappa_2$ (i.e., $\Psi(s)$ is a quadratic function), in the case of a Gaussian probability distribution, an important class of statistics we will discuss in the next section. They therefore allow one to test whether a given variable is Gaussian or close to being Gaussian. It should be noted that there are many distributions which lead to divergence of the power series (2.49), but that, even in such cases, the cumulants are defined by the (now formal) process described above, leading to perfectly definite formulas, (2.50), for the cumulants, always presuming convergence of the moments.

Extensions can be made to the case of multiple variables. For instance, with two variables,

$$\varphi(s_1, s_2) = \overline{e^{i(s_1U_1+s_2U_2)}} = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} e^{i(s_1U_1+s_2U_2)} P(U_1, U_2)\, dU_1 dU_2 \tag{2.51}$$

is the first characteristic function, the Fourier transform of $P(U_1, U_2)$. The expansion of $\varphi$ is

$$\varphi(s_1, s_2) = \sum_{n,m=0}^{\infty} a_{nm} \frac{(is_1)^n (is_2)^m}{n!m!} = e^{i(s_1\overline{U_1}+s_2\overline{U_2})} \sum_{n,m=0}^{\infty} \mu_{nm} \frac{(is_1)^n (is_2)^m}{n!m!} \tag{2.52}$$

where

$$a_{nm} = \overline{U_1^n U_2^m}, \qquad \mu_{nm} = \overline{u_1^n u_2^m} \tag{2.53}$$

are two-variable moments. For statistically independent variables

$$\varphi(s_1, s_2) = \overline{e^{is_1U_1} e^{is_2U_2}} = \overline{e^{is_1U_1}}\;\overline{e^{is_2U_2}} = \varphi_1(s_1)\varphi_2(s_2) \tag{2.54}$$

and, more generally, the characteristic function of any number of independent variables is the product of their characteristic functions. A second characteristic function is obtained by taking the logarithm of $\varphi(s_1, \ldots, s_N)$ and its (possibly formal) power series expansion yields the cumulants

$$\begin{aligned}
\kappa_1^{(i)} &= \overline{U_i} \\
\kappa_2^{(ij)} &= \overline{u_i u_j} \\
\kappa_3^{(ijk)} &= \overline{u_i u_j u_k} \\
\kappa_4^{(ijkl)} &= \overline{u_i u_j u_k u_l} - \overline{u_i u_j}\;\overline{u_k u_l} - \overline{u_i u_k}\;\overline{u_j u_l} - \overline{u_i u_l}\;\overline{u_j u_k}
\end{aligned} \tag{2.55}$$

as coefficients. For Gaussian variables, all cumulants above $\kappa_2^{(ij)}$ are zero, as for a single variable. Among other uses, this provides a basis for testing how near to joint Gaussian a given set of variables are.

## 2.3 Gaussian Statistics and the Central Limit Theorem

A single variable is said to be Gaussian (or normal) if $P(U)$ has the form

$$P(U) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(U-\overline{U})^2}{2\sigma^2}} \tag{2.56}$$

which is shown in Figure 2.6a and has the well-known symmetric bell shape with a single hump centered on $\overline{U}$. The mean, $\overline{U}$, and standard deviation, $\sigma$, are the only parameters defining the Gaussian distribution and, if we multiply a Gaussian variable by a constant or add a constant to it, the result remains Gaussian, with different mean and standard deviation. From (2.56) it can be shown that the skewness and flatness factors of a Gaussian variable are $S = 0$, $T = 3$, departures from which can be used as measures of lack of normality of a given variable (although one can have $S = 0$, $T = 3$ for a non-Gaussian distribution). Figures 2.6b and 2.6c illustrate distributions with $S \neq 0$ and $T \neq 3$. Two random variables are jointly Gaussian if

$$P(U_1, U_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\left\{-\frac{1}{2}\frac{1}{1-\rho^2}\left[\frac{u_1^2}{\sigma_1^2} + \frac{u_2^2}{\sigma_2^2} - 2\rho\frac{u_1 u_2}{\sigma_1\sigma_2}\right]\right\} \tag{2.57}$$

where $u_n = U_n - \overline{U_n}$, $\sigma_n$ is the standard deviations of $U_n$, and $\rho$ is the correlation coefficient of the two variables. Thus, in addition to the two means and standard deviations, the correlation coefficient enters as a parameter.

In general, an arbitrary number of random variables, $U_1, \ldots, U_N$, are said to have joint Gaussian (or normal) statistics if their joint probability distribution function has the form $\exp q(U_n)$, where $q(U_n)$ is a quadratic function of the $U_i$. The distribution function is then

$$[\det(2\pi R)]^{-1/2}\exp\left\{-\frac{1}{2}\mathbf{u}^T R^{-1}\mathbf{u}\right\} \tag{2.58}$$

where $\mathbf{u}$ represents the column vector formed from the $N$ fluctuations $u_n = U_n - \overline{U_n}$ and $R_{nm} = \overline{u_n u_m}$ is the positive-definite, symmetric matrix of correlations. Thus, the means and correlations of jointly Gaussian variables suffice to fix their probability distribution function and hence the full statistics of the variables. The distribution (2.58) has a single maximum at the mean value, $U_n = \overline{U_n}$, and drops off rapidly as $|U_n - \overline{U_n}|/\sigma_n$ increases, as in Figure 2.6a. It can be shown that the sum of jointly Gaussian variables is Gaussian, while if $U_1, \ldots, U_N$ are *independent* variables that are individually Gaussian, they are also jointly Gaussian, since the joint distribution of independent variables is the product of the individual distributions and the product of exponentials is the exponential of the sum.

Suppose that the time-dependent variable $U(t)$ is statistically steady and Gaussian. By saying that the process $U(t)$ is Gaussian, we mean that, no matter what $N$ is used, the values $U(t_1), \ldots, U(t_N)$ of $U(t)$ at any $N$ times are jointly Gaussian. Steadiness implies that $\overline{U}$ is independent of time and that the correlation matrix $R_{nm} = R(t_n - t_m)$, where $\overline{u(t+\tau)u(t)} = R(\tau)$ is the correlation function of $U(t)$. Thus, for steady Gaussian processes, giving $\overline{U}$ and $R(\tau)$ suffices to determine the full (i.e., $N$-time, for any $N$) statistics of $U(t)$. From $\overline{u(t+\tau)u(t)} = R(\tau)$, it is easily shown that $R(-\tau) = R(\tau)$, but considerably harder to demonstrate Kinchin's theorem that $R(\tau)$ is the Fourier transform of a *positive* function, known as the frequency spectrum. Provided $R(\tau)$ has these properties, it can also be shown that a statistically steady, Gaussian process, $U(t)$, can be constructed that has the given $R(\tau)$. Similar results hold for statistically steady vector functions of space and time, such as the velocity, $U_i(\mathbf{x}, t)$, in steady flow.

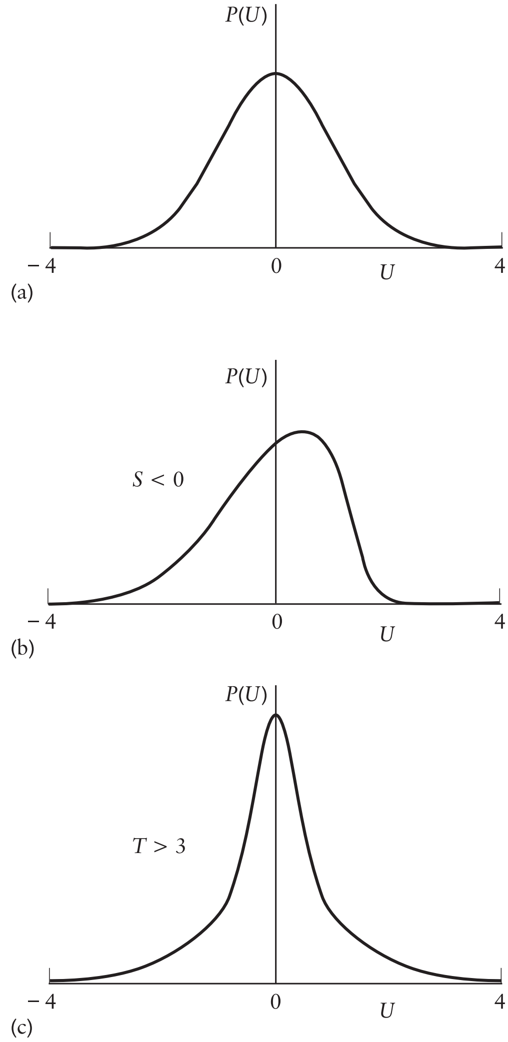


**Figure 2.6.** Comparison of (a) the normal distribution (for which $T = 3$, $S = 0$) with two other distributions, illustrating the cases (b) $S < 0$ and (c) $T > 3$. All distributions have zero mean and variance 1. The distribution in (b) is skewed, while that in (c) has more extensive tails than a Gaussian distribution, hence $T > 3$.

As noted in the previous section, the cumulants of a Gaussian distribution are zero at orders above two. This follows from taking the Fourier transform of (2.58) to obtain the first characteristic function. The result is the exponential of a quadratic function of the transform variables, so the second characteristic function yields that quadratic function and its series expansion terminates at order two. Setting the fourth line of (2.55) to zero yields

$$\overline{u_i u_j u_k u_l} = \overline{u_i u_j}\,\overline{u_k u_l} + \overline{u_i u_k}\,\overline{u_j u_l} + \overline{u_i u_l}\,\overline{u_j u_k} \tag{2.59}$$

which allows us to express fourth-order central moments in terms of second-order correlations for jointly Gaussian variables. This result forms the basis of the so-called quasi-normal approximation for closing the statistical equations of turbulence, in which the turbulent velocity fluctuations are assumed to be a sufficiently good approximation to Gaussian variables that (2.59) can be used to express their fourth-order moments.

The importance of Gaussian statistics derives from a profound result of the theory of probability which concerns sums of independent variables: the celebrated *central limit theorem*. Let $U_1, \ldots, U_N$, be statistically independent variables with identical distributions functions, then $U = U_1 + \cdots + U_N$ approaches a Gaussian distribution as $N \to \infty$. The requirement that the variables be identically distributed, present in this, the basic version of the theorem, can, in fact, be relaxed considerably.[4] The more important condition is that of independence and what the theorem says is that the sum of a *large* number of *independent* variables will be close to Gaussian. When the variables have different distributions one can determine the limiting distribution from (2.56) with

$$\overline{U} = \sum_{i=1}^{N} \overline{U_i}, \quad \sigma^2 = \sum_{i=1}^{N} \sigma_i^2 \tag{2.60}$$

[4] Proofs of the theorem require that the distribution functions of the random variables in the sum should satisfy certain mathematical conditions, the details of which we do not go into.

where a large number of terms must contribute significantly to the second sum, otherwise one can ignore the effects of all but the small number of variables which do, in which case there is no reason why $U$ should be close to Gaussian.

There is a potentially important, and not widely appreciated, restriction on the central limit theorem, which is that the asymptotic description as a Gaussian distribution, although accurate over the most probable part of the distribution of $U = U_1 + \cdots + U_N$, is generally invalid in the tails of that distribution. This non-uniformity occurs when $U - \overline{U}$ is of order $N$, much larger than the standard deviation, $O(N^{1/2})$, of $U$ and therefore representing rarely attained departures from the mean. Study of the asymptotic behavior when $U - \overline{U} = O(N)$ is known as large deviations theory (see, e.g., Varadhan (1984)), which has important applications to statistical mechanics and, more importantly for present purposes, to the statistical properties of the small scales of turbulence (see, e.g., Frisch (G 1995), section 8.6.4).

The reader may care for an example, demonstrating both the central limit theorem and its large deviation restriction. As assumed in the simplest version of the theorem, let $U_i$ be identically distributed, independent random variables and $U^{(N)} = U_1 + \cdots + U_N$ be the sum of the first $N$, so that $U^{(N+1)} = U^{(N)} + U_{N+1}$ gives $U^{(N+1)}$ as the sum of two independent variables. The reader can show that, if $V_1$ and $V_2$ are independent, with distributions $P_1(V_1)$ and $P_2(V_2)$, then the distribution of $V = V_1 + V_2$ is given by

$$P_+(V) = \int_{-\infty}^{\infty} P_1(V_1)P_2(V - V_1)dV_1 \tag{2.61}$$

which is the convolution integral of $P_1$ and $P_2$ (hint: owing to independence, the joint distribution of $V_1$, $V_2$ is $P(V_1, V_2) = P_1(V_1)P_2(V_2)$, while $P_+(V)dV$ is the integral of $P(V_1, V_2)$ over the infinitesimal strip $V < V_1 + V_2 < V + dV$ in the $(V_1, V_2)$ plane). That is, the distribution of a sum of independent variables is the convolution of their distributions. Applying this result with $V_1 = U^{(N)}$ and $V_2 = U_{N+1}$, we have

$$P^{(N+1)}(U) = \int_{-\infty}^{\infty} P^{(N)}(V)P(U - V)dV \tag{2.62}$$

where $P^{(N)}$ is the distribution of the sum $U^{(N)} = U_1 + \cdots + U_N$ and $P$ denotes the distribution function of the $U_i$. In other words, each time an extra term is added to the sum, its distribution is convolved with $P$, beginning with $P^{(1)}(U) = P(U)$. In the example we have in mind, the terms, $U_i$, in the sum have the Poisson distribution $P(U) = e^{-U}$ for $U > 0$, $P(U) = 0$ when $U < 0$, and the reader may verify that

$$P^{(N)}(U) = \frac{U^{N-1}e^{-U}}{(N-1)!} \tag{2.63}$$

for $U > 0$, $P^{(N)}(U) = 0$ when $U < 0$, reduces to $P(U)$ if $N = 1$ and satisfies the recurrence relation (2.62). The reader is encouraged to calculate and plot the probability distribution, $N^{1/2}P^{(N)}(N^{1/2}(V + N))$, of the normalized variable $V = (U - N)/N^{1/2}$ as a function of $V$ for a series of increasing values of $N$, using a computer (the normalization is based on the mean, $\overline{U^{(N)}} = N$, and standard deviation, $N^{1/2}$, of $U^{(N)}$). This bears out the central limit theorem graphically, with quite small $N$ sufficient for a roughly Gaussian-looking curve, although later convergence is rather slow.